

Fig 3a. CLUSTAL W (1.82) multiple sequence alignment

```

Mouse NsG33      HASAHASALLCALCCGLLAASAHAGYSEDRCSWRGSLTQEPGSVGQLTLDCTEGAIEWL
Rat NsG33       ---MLVAALLCALCCGLLAASARAGYSEDRCSWRGSLTQEPGSVGQLTLDCTEGAIEWL
Human NsG33      -MGFPAAALLCALCCGLLAPAAARAGYSEERCSSWRGSLTQEPGSVGQLALACAEGAVEWL
                  .:*****.:*:*****:*****:*****: * *:***:***

Mouse NsG33      YPAGALRLTLGGPDPGTRPSIVCLRPERPFAQAQVFAERMTGNLELLLAEGPDLAGGRCM
Rat NsG33       YPAGALRLTLGGSDPGTRPSIVCLRPTRPFAGAQVFAERMAGNLELLLAEGQGLAGGRCM
Human NsG33      YPAGALRLTLGGPDPRARPGIAICLRPVRPFAGAQVFAERAGGALLELLAEGPGPAGRCV
                  *****.*.*.*.*.* ***** * ***** . *****:

Mouse NsG33      RWGPRERRALFLQATPHRDISRRVAAFREFELHEDQRAEMSPQAQGLVDGACRPCSDAEL
Rat NsG33       RWGPRERRALFLQATPHRDISRRVAAFQFELHEDQRAEMSPQAQGFVDGACRPCSDAEL
Human NsG33      RWGPRERRALFLQATPHQDISRRVAAFREFELREDGRPELPPQAHGLVDGACRPCSDAEL
                  *****:*****:***:* *.*.:***:*:*****:*****

Mouse NsG33      LLAactsdfvIHGTiHGVaHdTeLQESvITVvVArViRqTLpLfKegSseGqGrASIRtL
Rat NsG33       LLtactSdfvIHGTiHGVvHdMeLQESvITVvAtRvIRqTLpLfQEGSseGRqQASvRtL
Human NsG33      LLAAcTSdFvIHGIiHgVtHdVeLQESvITVvAARvLRqTpPlFqAGrSGdQGLtSiRtP
                  **:***** *.*.* *****.:**:* * *: * * .*: *:*

Mouse NsG33      LRCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFsRVySAALtTHLNPCEmALD (SEQ ID NO:8)
Rat NsG33       LRCGVRPGPGSFLFMGWSRFGEAWLGCAPRFQEFsRVySAALAAHLNPCEVALD (SEQ ID NO:13)
Human NsG33      LRCGVHPGPGTFLFMGWSRFGEARLGCAPRFQEFRRAYEAARAAHLHPCEVALH (SEQ ID NO:3)
                  *****:*****:*****:*****: * *:*****:

```

Fig 3b. CLUSTAL W (1.82) multiple sequence alignment

Mouse	--MLVATLLCALCCGLLAASAHAGYSEDRCSWRGSGLTQEPGVSQGQLTLDCTEGAIEWLY	58
Rat	--MLVAALLCALCCGLLAASARAGYSEDRCSWRGSGLTQEPGVSQGQLTLDCTEGAIEWLY	58
Human	MGFPAALLCALCCGLLAPAAAGYSEERCSWRGSGLTQEPGVSQGLALACAEGAVEWLY ; .*:*****.:*:*****:*****:* *.***:*	60
Mouse	PAGALRLTLGGPDGPTRPSIVCLRPERPFAGAQQVFAERMGTGNLELLAEGPDLAGGRCMR	118
Rat	PAGALRLTLGGSDPGPTRPSIVCLRPTRPFAGAQQVFAERMAGNLELLAEGQGLAGGRCMR	118
Human	PAGALRLTLGGDPDRARPPIACLRPVRFAGAQQVFAERAGALELLAEGPGPAGGCVR *****.*.*.*.* ***** *	120
Mouse	WGPRERRALFLQATPHRDISRRAAFRFELHEDQRAEMSPQAQGLVDGACRPCSDAELL	178
Rat	WGPRERRALFLQATPHRDISRRVAAPFQELHEDQRAEMSPQAQGFVDGACRPCSDAELL	178
Human	WGPRERRALFLQATPHQDISRRVAAFRFELREDGRPELPQAHLGVGDGACRPCSDAELL *****:*****:***:* *.*:..***:*:*****	180
Mouse	LAACTSDFVIHGTHIGVHADTELQESVITVVVARVIRQTPLPLFKESGSEQGRASIRTLL	238
Rat	LTACTSDFVIHGTHIGVVHDMELQESVITTVATR VIRQTPLPLFQESSEG RGQASVRTLL	238
Human	LAACTSDFVIHGI IHGVTHDVELQESVITVVAARVL RQPPLFQAGRSGD QGLTSIRTPL *.***** ****.* *****.;**:* ***: * *.*:;* **	240
Mouse	RCGV RPGPSFLFMGW SRFG EAWLG CAP RFQEFSRVYS AALT THLN PC EM AL D	291 (SEQ ID NO:26)
Rat	RCGV RPGPSFLFMGW SRFG EAWLG CAP RFQEFSRVYS AALA AHL NP CE VAL D	291 (SEQ ID NO:13)
Human	RCGVHP GP GT FL FM GW SR FG EAR LG CA PR FQE FRAY EAARA AH LH PCE VAL H *****:*****:***** **.***:****.*:	293 (SEQ ID NO:3)

scoring matrix: BLOSUM50, gap penalties: -12/-2  
42.3% identity; Global alignment score: 747

```

      10      20      30      40      50 *
Innog. MRGAARAAWGRAGQPWPRPPAPGPPPPPLPLLLLLLAGLLGGAG-AQYSSDRCSWKGSGL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  -----MGFPAAALLCALCCGLLAPAAARAGYSEERCSSWRGSGL
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      10      20      30

      60      70      *      80      90      100      *110
Innog. THEAHRKEVEQVYLRCAAGAVEWMYPTGALIVNLR-PNTFSPARHLTVCI RSFTDSSGAN
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  TQEPGS--VGQLALACAEGAVEWLYPAGALRLTLGGPDPR--ARPGIACLRPVRPFAGAQ
      40      50      60      70      80      90

      120     130     140      *      150     160     170
Innog. IYLEKTG-ELRLVPDGDGRPRVQC--FG-LEQGGLFVEATPQQDIGRRTTG FQYELVR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  VFAERAGGALELLLAEGPG-PAGGRCVRWGP RRERRALFLQATPHQDISRRVAAFRFELRE
      100     110     120     130     140     150

      180     190     200     210     220
Innog. RHRAS--DLHEL SAP--CRPCSDTEVLLAVCTSDFAVRGSIQQVTHEPERQDS AIHLRV
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  DGRPELPPQAHGLVDGACRPCSDAELL LAACTSDFVIHGI IHGVTHDVELQESVITVVA
      160     170*      *      180      *      190     200     210

      230     240     250     260     270     280
Innog. SRLYRQKSRVFEPVPEGDGHWQG--RVRTLLECGVRPGHGDFLFTGHMHFGEARLGCAPR
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  ARVLRQT PPLFQAGRSGD---QGLTSIRTPLRCGVHPGPGTFLFMGWSRFGEARLGCAPR
      220     230     240      *      250     260      *

      290     300      * 310
Innog. FKDFQRM YRDAQERGLNPCEVGT D (SEQ ID NO:40)
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
NsG33  FQEFRRAYEAARA AHLHPCEVALH (SEQ ID NO:3)
      270     280     290

```

Fig 5

6/13

```
1  gcttcgccgg ggcggggcgg ccggcgcccc cggctgtccc cgcgcgccgc cggaccgcgc cccgcgcggg gcagcgggtg tgagagcccc gactccccgg
101 acgcgcgccg ccgtgccatg gggttccccc cgcgcggcgt gctctgcgcg ctgtgtgcgc gcctcctggc ccgggctgcc cgcgcggcgt actccgagga
>>.....CDS.....>
    m g f p a a a l l c a l c c g l l a p a a r a g y s e
201 gcgctgcagc tggaggggca ggcgcctcac ccaggagccc ggcagcgtgg ggcagctggc cctggcctgt gcgaggggcg cggttgagtg gctgtaccgc
>>.....CDS.....>
    e r c s w r g s g l t q e p g s v g q l a l a c a e g a v e w l y p
301 gctggggcgc tgcgcctgac cctggggcgc cccgatccca gagcgcggcc cggcatgcgc tgtctgcggc cggttgcggc cttcgcgggc gcccaggtct
>>.....CDS.....>
    a g a l r l t l g g p d p r a r p g i a c l r p v r p f a g a q v
401 tcgcggagcg cgcagggggc gccctggagc tgctgtggc cgagggcccc ggcgcggcag gggcgcctg cgtgcgctgg ggtccccgcg agcgccgggc
>>.....CDS.....>
    f a e r a g g a l e l l l a e g p g p a g g r c v r w g p r e r r
501 cctcttctcg caggccacgc cgcaccagga catcagccgc cgcgtggcgc ccttcgctt tgagctgcgc gaggacgggc gccccgagct gcccccgcag
>>.....CDS.....>
    a l f l q a t p h q d i s r r v a a f r f e l r e d g r p e l p p q
601 gccacggtc tcggcgtaga cggtgccctgc aggcctgca gcgacgtga gctgctcctg gccgcatgca ccaggcactt cgtaattcac gggatcatcc
>>.....CDS.....>
    a h g l g v d g a c r p c s d a e l l l a a c t s d f v i h g i i
701 atggggtcac ccatacgtg gagctgcagg agtctgtcat cactgtggtg gcgcgcctg tctccgca gacaccgcg cgtgtccagg cggggcgcgc
>>.....CDS.....>
    h g v t h d v e l q e s v i t v v a a r v l r q t p p l f q a g r
801 cggggaccag gggctgacct ccattcgtag cccactgac tgtggcgtcc acccgggcc acccaccctc ctcttcattg gctggagccg ctttggggag
>>.....CDS.....>
    s g d q g l t s i r t p l r c g v h p g p g t f l f m g w s r f g e
901 gcccggtgg gctgtgcccc acgattccag gagtccgc gtgcctacga ggctgccgt gctgcccacc tccaccctg cgaggtggcg ctgcactgag
>>.....CDS.....>
    a r l g c a p r f q e f r r a y e a a r a a h l h p c e v a l h
1001 gggctgggtg ctggggaggg gctggtagga gggagggtgg gccactgct ttggaggtga tgggactatc aataagaact ctgttcacgc aaaaaaaaaa
1101 aaaaaaaaa (SEQ ID NO:2)
```

Fig 6

1001 gggctgggtg ctggggaggg gctggtagga gggagggtgg gccactgct ttggaggtga tgggactatc aataagaact ctgttcacgc aaaaaaaaaa (SEQ ID NO:3)

```

1  ccacgcgtcc gccacacgct ccgcgcgtct ttgcgcgctc tttgcggcc tcttgccgc gtcgcgtcac gctgggtact cggaagacccg ctgcagctgg
>>.....CDS.....>
   h a s a h a s a l l c a l c c g l l a a s a h a g y s e d r c s w
101 aggggcagcg gtttgacca ggagcctggc agctgagccct ggactgtact gagggcgcta tcgagtggct gtaccacgct ggggcgctgc
>>.....CDS.....>
   r g s g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l
201 gctgaccct gggcgccccc gatccgggca cagcgccacag catcgtctgt ctgcgccacg agcgccctt cgctgggtgcc caggtctctcg ctgaacgtat
>>.....CDS.....>
   r l t l g g p d p g t r p s i v c l r p e r p f a g a q v f a e r
301 gacgggaat cttagattgc tactggccga gggcccgac ctggctggg gccgtgcat gcgtggggt ccccgcgagc gccgagccct ttctctgcag
>>.....CDS.....>
   m t g n l e l l a e g p d l a g g r c m r w g p r e r a l f l q
401 gccacaccac acccgacat cagccgcaga gttgctgcct tccgtttga actgcacgag gaccaacgtg cagaaatgtc tcccagggt caaggtcttg
>>.....CDS.....>
   a t p h r d i s r r v a a f r f e l h e d q r a e m s p q a q g l
501 gtgtggatgg tgctgcagg ccttgcagt atgcgagct cctcctggct gcatgcacca gtgattttgt gatccacggg accatccatg gggtcgcccc
>>.....CDS.....>
   g v d g a c r p c s d a e l l l a a c t s d f v l h g t i h g v a
601 tgacacagag ctgcaagaat cagtcacac tggtgtggtt gctcgtgtca tccgccagac actgcacactg ttcaagggaag ggagctcgga gggccaaagg
>>.....CDS.....>
   h d t e l q e s v i t v v a r v i r q t l p l f k e g s s e g q g
701 cgggcctcca ttctgtacctt gctgcgctgt ggtgtcgtc ctggcccagg ctccttcttc ttcatgggct ggagccgatt tggcgaaagt tggctgggct
>>.....CDS.....>
   r a s i r t l l r c g v r p g p g s f l f m g w s r f g e a w l g
801 gtgtccccg cttccaagag ttacgcgtg tctattcagc tgctctcag acccatctca acccatgtga gatggcactg gactgagaga cctggggagca
>>.....CDS.....>
   c a p r f q e f s r v y s a a l t t h l n p c e m a l d - (SEQ ID NO:8)
901 agccctggat ggaccttctt ctggagatgg ggtgttgggg aggtgtgatgg gaggttgggt gagaagggtg tggctcggat ggcatccttg taccacagct
>>.....CDS.....>
1001 gagctggtag aataactaagt aatctggacc ataaaaaaaa aaaaaaaa (SEQ ID NO:7)

```

Fig 7a

```

1  gggcagccgc gccgcgggct gctcgcgctg cggccccgac cctccccggc cagcagtcgc agggccccgc ggcgtcccta accatgctgg
   >>CDS.>
   m 1
91  tagccacgct tctttgcgc ctctgttgcg gcctcttggc cgcgtccgct cagcgtggct actcggaaga cgcctgcagc tggaggggca
   >.....CDS.....>
   v a t l l c a l c c g l l a a s a h a g y s e d r c s w r g
181  gcggtttgac ccaggagcct ggcagcgtgg ggcagctgac cctggactgt actgaggcgc ctatcgagtg gctgtacca gctggggcgc
   >.....CDS.....>
   s g l t q e p g s v g q l t l d c t e g a i e w l y p a g a
271  tgcgctgac cctggggcgc ccgcatcgc gcacacggcc cagcatcgtc tgtctgcgcc cagagggccc cttcgtggt gcccaggtct
   >.....CDS.....>
   l r l t l g g p d p g t r p s i v c l r p e r p f a g a q v
361  tcgctgaacg tatgaccgc aatctagagt tgctactggc cgaggggccc gacctggctg gggggcgcctg catgcgctg ggtccccgcg
   >.....CDS.....>
   f a e r m t g n l e l l l a e g p d l a g g r c m r w g p r
451  agcgcgagc ccttttcctg caggccacac cacacgcga catcagcgc agagttgctg ccttcggtt tgaactgcac gaggaccaac
   >.....CDS.....>
   e r a l f l q a t p h r d l s r r v a a f r f e l h e d q
541  gtgcagaaat gtctccccag gctcaaagtc ttggtgtgga tgggtcctgc agggcctgca gtgatgcga gtcctcctg gctgcacgca
   >.....CDS.....>
   r a e m s p q a q g l g v d g a c r p c s d a e l l l a a c
631  ccagtgattt tggatccac gggaccatcc atgggtcgc ccacacaca gagctgcaag aatcagtcac cactgtggtg gttgctcgtg
   >.....CDS.....>
   t s d f v i h g t l h g v a h d t e l q e s v i t v v a r
721  tcacccgcca gacactgcca ctgttcaagg aaggagctc ggaggggccaa ggccgggccc ccattcgtac cttgctgcgc tgtggtgtgc
   >.....CDS.....>
   v i r q t l p l f k e g s s e g q g r a s i r t l l r c g v
811  gtcctggccc aggtccttc ctcttcattg gctggagccg atttgcgaa gcttggtgg gctgtgctcc ccgcttccaa gagttcagcc
   >.....CDS.....>
   r p g p g s f l f m g w s r f g e a w l g c a p r f q e f s
901  gtgtctattc agtgtcttc acgaccatc tcaaccatg tgagatgga ctggactgag agacctggga gcaagccctg gatggacctt
   >.....CDS.....>
   r v y s a a l t t h l n p c e m a l d - (SEQ ID NO:26)
991  ctctcggaga tggggtgtg gggagggtga tgggaggtg ggtgagaagg gttggctcg gatggcatcc tggatccac agtgagctgg
1081  tagaatacta agtaactctg accataccag ccacttagt catggcttc tgggacgc agcatacca gctctgtgc tgcctcactt
1171  tgtctactct ccagtctgct gcccttctaa ccttcttag cctgctgacc agtgagctca tgtttcctc gaattccagg gtgctgctgg
1261  ggttcagagc aacgtgccc tagtttggaa gacttgagct aattgtttt tttttgttt ttttttaagg tggcctgggg
1351  gggggggcaa aca (SEQ ID NO:25)

```

Fig 7b

9/13

```

1  atgctggttag cggcgcttct ctgcgcgctg tgcgcggcc tcttggtgc gtccgtctga gctggctact ccgaggaccg ctgcagctgg aggggcagcg
>>.....CDS>
    m l v a a l l c a l c c g l l a a s a r a g y s e d r c s w r g s
101 gttgacca ggaacctggc agcgtgggc agctgacct ggattgtact gaggtgcta tcgagtggct gatatcagct ggggcgctgc gcctgactct
>.....CDS>
    g l t q e p g s v g q l t l d c t e g a i e w l y p a g a l r l t
201 aggcggctct gatccgggca cgcggccag catcgtctgt ctgcgccaa cacggccctt cgtcgtgccc caggtcttcg ctgaacggat ggccggcaac
>.....CDS>
    l g g s d p g t r p s i v c l r p t r p f a g a q v f a e r m a g n
301 cttagattgc tactggcca gggccaaggc ctggctggg gccgtgcat gcgtgggt cctgcgagc gccgagccct ttctctgcag gccacgccac
>.....CDS>
    l e l l l a e g q g l a g g r c m r w g p r e r a l f l q a t p
401 accgggacat cagccgcaga gttgctgcct tccaattga actgcacgag gaccaacgtg cagaaatgc tccccaggcc caagggtttg gtgtggatgg
>.....CDS>
    h r d i s r r v a a f q f e l h e d q r a e m s p q a q g f g v d
501 tgcctgcagg ccctgcagtg atgccgagct cctctgact gcattgcacca gtgactttgt gatccatggg accatccatg gggctcgtcca tgacatggag
>.....CDS>
    g a c r p c s d a e l l l t a c t s d f v i h g t i h g v v h d m e
601 ctgcaagaat cagtcacac tgtgtgtggc actcgtgtca tccgccagac actgccactg ttccaggaa gtagctcgga gggccggggc caggccctccg
>.....CDS>
    l q e s v i t v v a t r v i r q t l p l f q e g s s e g r g q a s
701 ttcgtacctt gttgcgtgt ggtgtgcgtc ctggcccagg ctcttctc ttcatgggt ggagccgatt tggcgaagct tggctgggt gcgctcccg
>.....CDS>
    v r t l l r c g v r p g p g s f l f m g w s r f g e a w l g c a p
801 cttccaagag ttcagccgtg tctattcagc tgctctcgc gccaccta acccatgta ggtggcactg gactgagaga cctggggagca agccctggat
>.....CDS>
    r f q e f s r v y s a a l a a h l n p c e v a l d - (SEQ ID NO:13)
901 ggatcttct ctgggggatgg ggtgttggg aggggtgata ggagggtggg tgggaagggt gtggctcaga tggcatcctg gtaccacag tgaggtggta

1001 gaatactaaa taacctggat cacacc (SEQ ID NO:12)

```

Fig 8